Metabolomics technology and bioinformatics for precision medicine

Rajeev K. Azad and Vladimir Shulaev

Corresponding author: Vladimir Shulaev, Department of Biological Sciences, BioDiscovery Institute, University of North Texas, Denton, TX 76210, USA. Tel.: 940-369-5368; Fax: 940-565-3821; E-mail: shulaev@unt.edu

Abstract

Precision medicine is rapidly emerging as a strategy to tailor medical treatment to a small group or even individual patients based on their genetics, environment and lifestyle. Precision medicine relies heavily on developments in systems biology and omics disciplines, including metabolomics. Combination of metabolomics with sophisticated bioinformatics analysis and mathematical modeling has an extreme power to provide a metabolic snapshot of the patient over the course of disease and treatment or classifying patients into subpopulations and subgroups requiring individual medical treatment. Although a powerful approach, metabolomics have certain limitations in technology and bioinformatics. We will review various aspects of metabolomics technology and bioinformatics, from data generation, bioinformatics analysis, data fusion and mathematical modeling to data management, in the context of precision medicine.

Key words: precision medicine; metabolomics; metabolite profiling; NMR; mass spectrometry; GC-MS; LC-MS; bioinformatics

Introduction

Precision medicine is rapidly emerging as a strategy to tailor medical treatment to a small group or even individual patients that have a similar genetic background, environment and lifestyle and, therefore, similar treatment prognosis. It has been shown that ethnic and gender differences can account for variation in responses to treatment. In 2016, the White House announced the precision medicine initiative (PMI), a long-term research program involving the National Institutes of Health (NIH) and multiple other research centers, which aims to 'enable a new era of medicine through research, technology, and policies that empower patients, researchers, and providers to work together toward development of individualized care' (https://obamawhitehouse.archives.gov/node/333101). Precision medicine relies on progress in systems biology and omics disciplines, including genomics, transcriptomics, proteomics and metabolomics.

Metabolomics is a global approach that can provide measurements of all, or a large number of metabolites in cells, tissues or biological fluids. Major approaches used in metabolomics studies include targeted analysis, metabolite profiling and metabolic fingerprinting [1, 2]. As a global approach that can measure all or a large number of cellular metabolites, metabolomics is distinctively positioned to provide a unique metabolic readout of patient's physiological or disease state [3]. Mass spectrometry (MS)-based and nuclear magnetic resonance (NMR)-based approaches are now routinely used for newborn screening [4–8], drug screening [9, 10], pharmacometabolomics [11–18], bacterial identification [19–21], metabolic imaging [22–25] or gut flora analysis [26–29].

Rajeev K. Azad is an Associate Professor at the University of North Texas. His research interests are in the area of Bioinformatics and Computational Biology, particularly, development and application of mathematical and computational methods to understand how organisms, specifically microbes, innovate to adapt to changes in environment, study of large omics data sets to determine how organisms respond to stress at the molecular and physiological level and development of novel approaches to decipher structural and functional features in genomes and elucidate their relationships in the context of evolution.

Vladimir Shulaev is a Professor at the University of North Texas. He heads the Biochemical Profiling Group, which focuses on developing methods for high-throughput metabolite profiling and application of metabolomics to systems biology and to study stress response in microorganisms, plants and animals.

Several recent reviews highlighted the power of metabolomics as a tool for precision medicine [30–35]. A recent metabolomics community white paper stresses the importance of metabolomics for precision medicine and outlines major approaches for including metabolomics data in large PMIs [31]. Combination of metabolomics with sophisticated bioinformatics analysis and mathematical modeling can provide a metabolic snapshot of the patient over the course of disease and treatment or classify patients into subpopulations and subgroups requiring individual medical intervention.

Here, we will review various aspects of metabolomics technology and bioinformatics in the context of precision medicine.

Precision medicine and metabolic phenotypes

Precision medicine aims at developing the best treatment or prevention method based on a person's genetics, environment and lifestyle. Precision medicine can be broadly divided into precision prevention and precision treatment [36]. Prevention treatment encompasses an array of measures focused on assessing risk of individual toward particular disease and devise a preventive intervention to decrease this risk and prevent the development of disease [37]. One of the earliest examples of precision prevention is a newborn screening for genetic mutations, such as phenylketonuria, and development of dietary intervention to prevent the onset of disease. Precision treatment, on the other hand, aims at developing personalized treatment strategy to cure disease while taking into account patient's genetics, environment and lifestyle as well as individual treatment response parameters to predict posttreatment outcome.

Metabotypes and precision medicine

Metabolic phenotype or metabotype (also called chemotype) concept is often used to describe a particular metabolic state of biological system through its metabolic signature represented as different level of individual metabolites [38]. Metabotypes can be defined as absence or presence of a particular metabolites, absolute or relative concentration of metabolites in a sample or as metabolite profiles or metabolic signatures. Metabolic phenotyping, or metabotyping, can be best performed using various metabolomics approaches, from targeted analysis to metabolic fingerprinting [31, 39]. Multiple examples exist in the literature describing successful metabolic phenotyping or metabotyping using metabolomics (reviewed by [40–45]). Metabolomics aided in characterizing many disease-associated metabotypes in cancer [46, 47], inflammatory bowel disease (IBD) [48, 49], asthma [50], diabetes [51], traumatic brain injury (TBI) [52], metabolic syndrome [53], Parkinson's disease [54] and other pathologies. For example, Calvani and colleagues [55] used NMR-based metabolomics analysis to identify an obesity-associated metabotype that differs from that of lean controls, while Tam and colleagues [56] used liquid chromatography–mass spectrometry (LC-MS)-based/MS-based metabolic phenotyping to identify metabolic biomarkers in older people with late-onset type 2 diabetes mellitus. Precision medicine would require the development and characterization of metabotypes for individual patients and their response to treatment.

It was suggested that local microenvironmental cues likely contribute to disease progression and resolution in many diseases including cancer [57, 58], IBDs [59], liver disease [60], Dupuytren's contracture [61], fibrotic disease [62], autoimmune disease [63] and obesity [64]. For example, there are evidence

that many tumors are not homogeneous and exhibit different genetic and metabolic state. Hu and colleagues [57] have demonstrated heterogeneity of tumor-induced metabolic gene expression across 22 different cancer types. Metabolic heterogeneity also exists within an individual tumor tissue, which may require specialized treatments to target different tumor regions. Recently, Okegawa and coauthors [65] demonstrated intratumor heterogeneity in primary kidney cancer [58]. Using global metabolomics analysis of multiple spatially separated samples within tumors, they have shown that different portions of a human primary kidney tumor possess different metabolic characteristics and drug sensitivity [65]. Hensley *et al.* [58] demonstrated metabolic heterogeneity in human lung cancer. Metabolic heterogeneity within tumor may be attributed to heterogeneity in gene expression [66]. Extensive intratumor genetic heterogeneity was shown for spatially separated portions of primary renal carcinomas and associated metastatic sites [67]. In other examples, it was demonstrated that microenvironment of adipose tissue influences cardiovascular disorders, including atherosclerosis and ischemic heart diseases (reviewed by [64]), and local intestine microenvironment can significantly contribute to both disease progression and resolution in IBDs [59]. Novel metabolomics technologies can assist in understanding these metabolic tissue microenvironments by performing metabolomics analysis of small cell population or even single cells [68]. Precision medicine has to address this tissue heterogeneity and develop treatment for a specific metabolic microenvironment within diseased tissues [69].

Metabolomics technology

Human metabolome

It is important to define the complete composition of metabolites present in human organism, in other words the human metabolome, and relate it to the metabolome of an individual. Despite significant developments in metabolomics, we still do not have a full understanding of the complete human metabolome, which is highly complex and includes both endogenous and exogenous metabolites. Endogenous metabolites are naturally produced by human organisms, while exogenous metabolites, or xenobiotics, are not naturally produced by an organism and include chemicals to which humans are exposed over the course their life. Metabolome is influenced by many factors, including genetic background, diet, environmental exposures and gut microflora. In addition to genetic background, diet significantly contributes to the metabolome composition. For example, it is a source of essential amino acids or many plant-derived metabolites, such as flavonoids or polyphenols. Subsets of the metabolome, i.e. tissue-specific or disease-specific metabolomes, can be useful when treating specific organ or disease. Also, there is an effort to compile a metabolome linked to a specific disease, i.e. cancer metabolome, diabetes metabolome, etc. Human microbiome metabolome, or the metabolome of human gut microflora, can also provide an invaluable information for diagnostics and treatment [70]. We will examine the role of microbiome in the precision medicine later in the article while discussing meta-genomics and meta-metabolomics. The exposome, which encompasses total human exposure from environmental sources, was also implicated in many diseases [71–75]. Several projects and databases are focused on defining human metabolome as well as species-centered or disease-centered metabolomes, microbiome metabolome and exposome [27, 28, 76–83].

Sample collection and processing

The nature of biological sample, collection procedure and extraction methodology is a critical component of any metabolomics study. Typically, human metabolomics studies analyze either biological fluids or tissue samples. Biofluids are generally easier to collect and analyze. Biofluids used for metabolomics analysis include serum [84–86], plasma [11, 87], urine, cerebrospinal fluid, ascitic fluid [88], saliva [89–91], tear [92], bronchial wash (BW) and bronchoalveolar lavage fluid [93], seminal fluid [94], prostatic secretions [94] or fecal samples [95]. Tissue analysis is more complicated because of tissue heterogeneity and often limited sample volume. Recently, dried blood spots (DBSs) samples became popular for various clinical applications. DBS samples were successfully used in newborn screening [96–98] and pharmacological analyses [99–102] and are now being investigated in metabolomics analysis [103–105]. Although several standardized procedures were published for metabolomics analysis of various sample types (reviewed by [105–108]), more standardization on sample collection and processing is necessary.

It is important that all the required metadata, including age, gender, diet, physical activities, disease history, medications, etc., be collected and recorded during sample collection. Ideally, these factors that have dramatic influence on the metabotype should be taken into account during patient interview and treatment design and before the samples are collected.

Biobanking

Precision medicine would require the development of specialized systems for collecting and storage of biological specimens [22, 31, 109–111]. Samples have to be collected under standardized conditions to ensure reproducibility and critical metabolites preservation. Abuja *et al.* [110] performed systematic study on the influence of storage scenarios on the liver metabolome with different storage temperatures and repeated transfer of samples between storage and retrieval environments, which simulates the typical biobanking conditions. They have shown that storage temperature affected metabolite concentrations only little, while the number of temperature change cycles has a strong effect on metabolites stability [110]. Biobanking is considered as a critical component of precision medicine workflow and is being incorporated into several PMIs [112–114]. For example, MOBIT project adopted the oncology biobanking procedures developed by a commercial biotech company Indivumed GmbH (Hamburg, Germany) [22]. The workflow incorporates several steps that are important for metabolomics analysis and consists of following major steps: patient qualification into biobank, anonymization of patient data, collection of biospesimens and patient data, preparation and storage of samples into biobank and quality control of tissues and biofluids. Another example of successful biorepository infrastructure is an open-source biorepository management system developed by Felmeister and coworkers [114]. The system consists of electronic Honest Broker (eHB) and Biorepository Portal (BRP) that, in tandem, allow for integration of clinical, specimen and genomic data collected for biorepository resources while protecting patient privacy [114]. As of January 2016, eight institutions were participating in biobanking activities using this tool kit with over 4000 unique subject records deposited in the eHB and over 30 000 specimens accessioned [114].

Metabolomics approaches

Metabolomics technology has significantly evolved over the past decade, and new techniques and approaches specifically suited for various clinical applications have been developed. Owing to chemical complexity of the metabolome, no single analytical methodology can measure all metabolites present in a biological sample; therefore, a combination of analytical techniques is used for metabolome analysis (reviewed by [1, 3]).

The major approaches used in metabolomics studies include targeted analysis, untargeted metabolite profiling and metabolic fingerprinting [1, 115]. Targeted analysis is a quantitative approach that measures the concentration of a limited set of known metabolites. Untargeted or global metabolomics attempts to measure large set of metabolites in the sample without knowing a priori which metabolites are expected to be present in each individual sample. Untargeted metabolite profiles have to be as comprehensible as possible to uncover the underlying metabolic network changed in a particular physiological or pathological state [116]. The biggest drawback of the untargeted profiling is that the majority of peaks in the profile are not identifiable. Application of high-resolution MS can significantly improve unknown compound identification, but accurate mass measurement alone is not sufficient for positive structure identification. Additional orthogonal information, such as chromatographic retention time, isotope pattern matching, multiple stage MS and collisional cross-section, is often needed for correct structural annotation [117, 118]. Metabolic fingerprinting considers a total metabolite profile as a unique pattern, or fingerprint, for a particular metabolic state without attempting to identify each individual metabolite in the profile. All three metabolomics approaches can be used in precision medicine.

Analytical techniques

Major analytical techniques currently used in metabolomics include NMR, Fourier transform infrared spectroscopy, gas chromatography–mass spectrometry (GC-MS), LC-MS and capillary electrophoresis–mass spectrometry. Precision medicine poses additional requirements for technology: ease of use, increased throughput, automation and real-time data generation and analysis. The advantages and drawbacks of different technologies used in metabolomics have been widely discussed [1, 3]; therefore, we will largely focus on techniques and instrumentation that are best suited for precision medicine.

NMR is widely used in metabolomics studies because of its nondestructive nature and ability to simultaneously measure many organic compounds present in the biological sample [3]. Being a nondestructive methodology makes NMR advantageous over other method, as the samples can be used for multiple assays. This is especially important when the sample size is limited, which is common in biomedical applications. The major limitation of the NMR as a comprehensive technique for metabolomics is its low sensitivity, which permits the detection of metabolites only at the micromolar level. Depending on the nature of the sample and specific biological questions, several NMR techniques can be used in precision medicine. ^{1}H-NMR is the most often used technique, although ^{31}P-NMR and ^{13}C-NMR can be applied as well. High-resolution magic angle spinning is useful when studying liquid or intact solid tissue samples [119–121]. NMR has been largely used for metabolic fingerprinting, biomarker discovery and metabolic flux analysis. It was extensively applied for metabolomics analysis and discrimination of numerous diseases, including cancer [122–124], kidney

diseases [125] sepsis [32, 126], cardiovascular diseases [127–129], acute lung diseases [130], Alzheimer's disease [89] and neonatology [131]. Different cancers, including prostate [132, 133], esophageal [134–136], breast [137], gastric [138], colorectal [123] [139], lung [140, 141], brain [122], ovarian [142] and renal [143], can be classified using NMR profiles. It was also successfully used to characterize metastatic cancer in lymph nodes [144].

MS-based methods provide increased sensitivity and ability to assay diverse range of cellular metabolites in wide polarity range. MS-based metabolite profiling can be performed either using shotgun approach based on direct infusion mass spectrometry or in combination with front-end separation technique (hyphenated techniques). Most often MS is combined with GC, LC, ion chromatography or capillary electrophoresis (CE) (reviewed by [1, 145, 146]). MS-based hyphenated techniques are extensively used for metabotyping of various diseases to characterize disease-associated metabolome. For example, Wei *et al.* [147] used GC-MS-based urine metabolomics combined with the personalized diagnosis guided by Chinese medicine (CM) to identify diagnostic biomarkers for prediabetic subtypes. In this study, three CM physicians reached 85% diagnosis consistency resulting in the classification of three prediabetic groups. The urine metabolic patterns of all three groups were clearly discriminated [147]. In another study, Jacobs *et al.* [48] used Ultra-Performance Liquid Chromatography - mass spectrometry (UPLC-MS) to characterize a disease-associated metabolomics state in relatives of IBD patients and demonstrated that healthy first-degree relatives can have dysbiosis associated with an altered intestinal metabolome. Similarly, Liu *et al.* [148] used capillary electrophoresis-time-of-flight mass spectrometry-based metabotyping to systemically study the differences between hemodialysis and high-flux hemodialysis on plasma metabolite changes in patients.

Currently, medical application of MS is dominated by triple quadrupole MSs, which are generally used for analysis of specific metabolite classes, such as amino acids [149], neurotransmitters [150], steroids [151] and drugs and their metabolites [152, 153]. With progress in metabolomics technology, which demands much broader metabolite coverage, other types of mass detectors are becoming more popular. Substantial advancements in metabolite profiling were achieved over the past decade as a result of a wide adoption of high-resolution mass spectrometry (HRMS), which offer several advantages over lower-resolution instruments [145, 146, 154, 155]. High mass accuracy and resolving power of the HRMS instruments enable better characterization of unknown metabolites by assignment of elemental formula and also allows for identification of adducts with high precision [155]. A variety of multiple ion sources are available to use with HRMS. Mass analyzers that can perform HRMS include Fourier transform ion cyclotron resonance (FT-ICR), Orbitrap and time-of-flight analyzers.

Insufficient sample volume, especially when analyzing tissue biopsies, which can contain as little as several thousand cells, often imposes a significant limitation on metabolomics technology. Microbore UPLC or capillary LC or LC-MS can be used to increase the sensitivity of the metabolomics analysis. There are significant developments in this area over the past years. Gray *et al.* [156] using 1 mm I.D. columns instead of 2.1 mm I.D. columns for metabolic phenotyping using UPLC-MS achieved equivalent or superior performance in terms of peak capacity and sensitivity. The increase in sensitivity of this method allows for using smaller sample volume [156]. Capillary LC-MS can be even more sensitive. Ni and coauthors [157] used a 75 μm inner diameter column coupled to a quadrupole ion

trap MS operated in full scan mode to analyze single islets of Langerhans, which are microorgans found in the pancreas that contain a few thousand cells each. Authors achieved high sensitivity with detection limits of 0.1–33 fmol for polar anions in 15 nl injection volumes. It was shown that number of detected metabolites in small samples can be achieved by reducing column diameter [158]. Simply reducing column I.D. from 50 to 25 μm, Edwards *et al.* [158] were able to increase the number of detected metabolites from 111 $(+/-9)$ to 156 $(+/-17)$ in *Escherichia coli* lysate samples. This improvement was attributed to an increase in separation efficiency, an increase in sensitivity and a decrease in adduct formation.

One of the limitations of the GC-MS- and LC-MS-based techniques is their inability to provide information on metabolites distribution within tissue or organ. This information is extremely critical for diagnostics and treatment of many diseases. The mass spectrometry imaging (MSI) can provide such an information. MSI can be performed using either vacuum ionization technique, such as vacuum matrix-assisted laser desorption/ionization (MALDI) [159–164], or ambient ionization method, such as atmospheric pressure MALDI (AP-MALDI) [165, 166], desorption electrospray ionization [167–170] and matrix-assisted laser desorption electrospray ionization [171, 172]. Most often MSI is performed using HRMS. MSI has been successfully applied for imaging of various human and animal tissues, such as brain [162, 173–176], heart [159–161, 177], liver [178], kidney [179], skin [180], breast [181–185] and lens [163, 164]. It has also been used to study clinical samples from cancer patients [181, 182, 184–186]. MSI can be performed on either cryosections [187] or formalin-fixed, paraffin-embedded (FFPE) human tissue samples [186], which provides the ability to reanalyze previously stored samples. For example, Buck and colleagues [186] analyzed FFPE tissues from 350 different cancer patients using high-resolution matrix-assisted laser desorption/ionization Fourier-transform ion cyclotron resonance mass spectrometry imaging (MALDI-FT-ICR MSI) and were able to discriminate between normal and tumor tissues, and different tumors from the same organ. They also found an independent prognostic factor for patient survival using MSI approach.

Rapid evaporative ionization mass spectrometry (REIMS) can go beyond other MSI applications [188–193]. It uses the aerosol by-product of electrosurgical (Bovie) tools [194]. St John and colleagues [194] developed an MS method for the rapid analysis of heterogeneous breast tissues, namely, iKnife device based on REIMS technology. They have shown that the iKnife is capable of accurately separating breast tissue types by interpretation of the cellular chemical constituents. Critical component of this technology is recognition software that enables real-time analysis of both *ex vivo* and *in vivo* breast tissue.

Bioinformatics for precision medicine

Bioinformatics in precision medicine is driven by the need to integrate omics biomarkers identified in individuals or cohorts of patients. These biomarkers could be single nucleotide polymorphisms (SNPs), structural variants, circulating DNAs, methylated DNAs, mRNAs and microRNAs (miRNAs), proteins and metabolites. Bioinformatics for precision medicine needs to support raw data generation and real-time processing, computational analysis, data and results visualization, data fusion, mathematical modeling, clinical data integration and data management. As per the National Research Council's guidelines, multilayer molecular or omics data, including exposome, genome, epigenome, transcriptome, proteome, metabolome

and microbiome, collected from individuals or cohorts of patients, should be deposited in common data repository, namely, Information Commons. In addition, clinical information and epidemiological data where available should be included to enable a comprehensive investigation of links between the data layers leading to a knowledge network. The knowledge network will enable taxonomic classification of diseases, rendering precision diagnosis and therapeutic intervention [195].

Metabolomics, similar to other 'omics' disciplines, requires specialized bioinformatics tools (reviewed by [1, 196–199]). Metabolomics data analysis involves several major steps and requires extensive raw data preprocessing followed by multivariate statistical analysis using specialized mathematical, statistical and bioinformatics tools, data mining and integration with other omics data and mathematical network modeling [196, 199–201]. In a recent review, Duerr-Specht et al. [202] defined three major areas in modern medicine and health care where improvement in health-care informatics and data management should focus on: organizational (including administrative and political), technological and educational. This is true for bioinformatics for precision medicine metabolomics. Precision medicine imposes special requirements on metabolomics data handling and bioinformatics analysis. Ultimately, metabolomics experiments focused on precision medicine will generate vast amount of raw data that have to be processed and analyzed, often in real time. Specific bioinformatics data processing and statistical tools used in metabolomics studies have been extensively reviewed [1, 154, 196, 203]; therefore, we will outline major strategies to analyze metabolomics data and largely focus on specific aspects of metabolomics and other omics bioinformatics relevant to precision medicine.

Data collection and processing

Translating the molecular data and statistical classifiers to useful products in clinic is a significant challenge, beginning from data procurement to clinical trial. Raw data almost always need to be preprocessed, which includes quality processing (e.g. data trimming), normalization or rescaling. In genomics and transcriptomics, the raw sequence reads are quality checked using programs, such as FASTQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/), and reads are trimmed to eliminate the low-quality nucleotides, often at the left and right ends of reads, using programs, such as Trimmomatic [204]. In metabolomics, raw data processing requires noise subtraction, chromatographic peak alignment, data normalization and scaling before the statistical analysis. Raw data preprocessing can be performed using instrument vendor-specific or open-source tools [205].

Omics data analysis

Early patient stratification method was based on genetic variation. Genomic DNA assays have been developed to interrogate DNAs in the samples of patients to identify nucleotide variants, including indels, i.e. insertions or deletions of one to few nucleotides and structural variants that encompass large genomic regions. These include targeted panel sequencing, whole-exome sequencing (WES) and whole-genome sequencing (WGS) [206]. Diana et al. [207] developed a targeted panel, 'The PulmoGene Test', which includes 64 genes implicated in lung disease that enabled interrogation of lung tissue samples for these genes. In contrast, WES and WGS entail sequencing of coding regions of all genes and the entire genome, respectively, to catalog single-nucleotide variants, indels and structural variants in genes or any genomic regions. While the latter is both expensive and time-consuming, the rapid decline in cost and turnaround time in recent years has raised the hope to use WES and WGS for the masses in precision medicine. The first step in identification of variants is alignment of sequence reads onto the reference human reference genome. This can be accomplished using publicly available aligners, such as BWA [208], or by using a toolkit, such as Genome Analysis ToolKit (GATK) (https://software.broadinstitute.org/gatk/) that also outputs SNPs and short indels. A number of variant calling programs have been developed, e.g. SIFT and PolyPhen for identifying pathogenic mutations, MoDIL for small indel detection [209], BreakDancer for structural variant detection [210] and XHMM for copy number variant detection from WES [211]. Programs, such as Cufflinks [212] or DESeq [213], can be used for evaluating the differential gene expression analysis once the RNA sequencing reads are mapped on the reference genome. Expression data in combination with variant data allow identification of variants that impact transcription and splicing.

Molecular classifiers have been developed to discriminate samples from patients and normal individuals or for disease stratification; however, these classifiers need to address issues with multiple comparisons, dimensionality problems [large number of features (p), small number of samples (n) problem] and validations. Molecular classifiers are becoming increasing popular in oncology because of the pressing need to differentiate cancers or cancer subtypes [214]. In addition to overcoming the technical challenges associated with molecular classifiers, Ferte et al. [215] suggest dynamic sampling of tumors, both on spatial and temporal scales, for an unbiased assessment that could be potentially rewarding in efforts to usher in the molecular technologies to clinics. Metabolic phenotyping has been used for static and dynamic patient stratification (reviewed by [216]). Unlike static patient stratification, which can predict the development of disease based on molecular classifiers, dynamic stratification can model the physiological and molecular responses to treatment and predict posttreatment outcomes [217–220]. For example, in a recent study, Elmariah et al. [221] performed LC-MS-based metabolite profiling of plasma from patients undergoing transcatheter aortic valve replacement (TAVR) to predict acute kidney injury (AKI) occurrence using previously identified plasma metabolites predictive of incident chronic kidney disease. They have shown that in elderly population with severe aortic stenosis undergoing TAVR metabolite profiling improves the prediction of AKI. Of 44 patients undergoing TAVR, 22 had chronic kidney disease and 9 developed AKI. Of 85 metabolites profiled, 5-adenosylhomocysteine successfully predicted AKI after TAVR and was also predictive of all-cause mortality after TAVR. Other examples of metabolic phenotyping-based patient stratification include prediction of exercise-induced ischemia in patients with suspected coronary artery disease [222], prediction of response to neoadjuvant chemotherapy for breast cancer using serum metabolite profiling [223] or prediction of the outcome of first systematic transrectal prostate biopsy using serum metabolomics [224].

Advances in machine learning techniques and artificial intelligence (AI) hold the promise to catalyze deep analysis and interpretation of big data in medicine, and facilitate integration of molecular and clinical data for accurate diagnosis and precision therapy. Although machine learning—often considered as an application of AI—has a history of several decades, the ushering in of the big data era has made these techniques in the

broad domain of AI even more relevant and sought after for decoding mysteries underneath the ocean of omics and other clinical data. Machine learning has been mainly categorized into supervised and unsupervised, the former requiring labeled data (training data) to let the model learn the 'known' patterns underlying the labeled data and then discern these patterns in yet unseen data (test data); in contrast, the unsupervised learning does not require labeled data and is by design endowed with the ability of discern unknown underlying patterns in an unbiased manner. Examples of supervised learning methods include neural networks, support vector machines, elastic net, decision trees, random forests, hidden Markov model, etc., although some of these methods can be adapted to unsupervised or semi-supervised learning. Examples of unsupervised methods include clustering algorithms, such as hierarchical agglomerative clustering, *k*-means clustering, principal component analysis (PCA) and self-organizing map. AI or machine learning has an unbounded potential in big data analysis, and in particular, in precision medicine where this enabling technology is at the core of 'deep data integration' (e.g. integration of gene expression data with transcription factor binding data to decipher genomic regularity network) [225, 226] and 'broad data integration' [226] (e.g. integration of different omics data, as well as clinical data, to gain a systems level understanding of a condition or disease). Recent studies have highlighted the importance of machine learning methods in diagnosis of complex diseases based on omics data, e.g. Zheng *et al.* [84] applied self-organizing maps to serum metabolomics data from renal cell carcinoma patients to identify biomarker metabolites for early diagnosis of renal cell carcinoma; Nielson *et al.* [227] reported successful application of topological data analysis, a machine learning tool, to imaging, genetics and clinical data from TBI patients for patient stratification and detection of candidate recovery biomarkers; machine learning by tensor factorization was used in precision medicine for heart failure with preserved ejection fraction, specifically in devising targeted therapies via disease subtyping based on phenotypic and omics data [228]; and to cite just a few among many applications in precision medicine.

These multivariate statistics tools including the unsupervised and supervised algorithms are commonly used to analyze metabolomics data [1, 196, 200, 229, 230]. Unsupervised methods that have been used in analyzing metabolomics data are hierarchical clustering [231], PCA [232] and self-organizing maps [233]. Supervised methods include support vector machines [234, 235], partial least squares ([236], 2003), analysis of variance [230], *k*-nearest neighbors and discriminant function analysis [237].

Data integration and fusion

Integrating different layers of molecular data, epidemiological data, medical image data and clinical data, for an individual or a cohort of individuals, or a population, and visualization are a pressing need. G-DOC Plus was developed recently to integrate big data, including omics and clinical data, and provide bioinformatics tools for analysis and interpretation, and generate hypotheses for further biomedical research [238]. The use of eHB and BRP open-source project is another effort toward data integration and collaborative research facilitated by anonymization of the specimens, thus bypassing the time-consuming institutional reviews and permission [114]. Pathology Integromics in Cancer (PICan) is a platform for integration of multimodal data and analysis for discovery of novel biomarkers

and their validation [239]. IMPACT integrates WES profiles with therapeutics. IMPACT detects the coding variants and predicts drugs that can neutralize the effects of the deleterious variants [240]. Associating phenotypes (e.g. metabolomics data that most closely represent the phenotypes) to genotypes (e.g. Single Nucleotide Polymorphism (SNP), Copy Number Variation (CNV) and expression data) is a key to the success of data integration or fusion but poses a significant challenge. Indeed, an important problem in precision medicine to associate the metabolic changes to the underlying genomic or transcriptomic changes in patients. Understanding variations in metabolome in the context of genomic alterations and gene expression changes is nontrivial, as these variations may be arising as a consequence of yet unknown gene–gene interactions or gene–environment interactions. In addition, posttranscriptional and posttranslational modifications may be players in rendering complex phenotypes, and therefore, deciphering these molecular modifications and molecular interactions in the cell (interactome) is critical to the translational advances in precision medicine [241].

Open-source software for metabolomics

Computational analysis of metabolic phenotypes for precision medicine can be facilitated by further development of open-source software tools for data processing, multivariate statistical analysis and machine learning, data integration and visualization as well as reference databases [242, 243]. Open-source software includes XCMS [244, 245] and XCMS Online [246], MZmine [247], xMSanalyzer [248] and OpenMS [249]. In combination with data analysis and multiple platform data integration tools, such as MetaboAnalyst [250–252], Metabolite Set Enrichment Analysis (MSEA)[253, 254], Molecular Networking Approach [242, 255] and Metabolic Pathway Analysis (MetPA) [256], these open-source software tools have been successfully used in metabolomics-driven precision medicine studies for biomarker discovery and validation as well as for patient stratification. Mass spectral libraries and publically available reference biochemical and pathway databases are designed to facilitate metabolite identification and data integration in metabolomics. There are many such databases available for precision medicine metabotyping most commonly used of which are METLIN [257–259], Kyoto Encyclopedia of Genes and Genomes (KEGG) [260], MassTRIX [261, 262], Madison Metabolomics Consortium Database (MMCD) [263], Human Metabolome Database and drug bank [79–82, 264], LIPID MAPS [265–267]. PubChem [268] and ChemSpider [269].

Precision metagenomics and meta-metabolomics in precision medicine

A new dimension in precision medicine now is the human microbiome. Understanding the host–microbiome interactions is central to understanding a number of diseases that are a consequence of dysbiosis including cancer [20, 27]. Microbiome sequencing has facilitated significant research in this area and has enabled development of microbiota based therapy [270]. Metagenome sequencing has made possible rapid detection of pathogens and understanding of the evolution of hypervirulent and multidrug-resistant strains. 'Precision Metagenomics' is a pipeline that leverages sequencing and bioinformatics to profile microbes in a sample, identify antimicrobial resistance markers and perform functional analysis to detect pathways that are enriched in a sample [271]. Recent effort in metabolomics analysis of gut microflora adds a vast volume of metabolomics data

that have to be integrated with metagenomics, metatranscriptomics and, eventually, clinical data [55, 70, 272–278]. This data integration would require the development of specialized bioinformatics pipelines capable of integrating data from multiple omics platform and incorporating data from large medical databases [279].

Initiatives on real-time tests/clinical trials based on omics data

Clinics and medical research centers are now embracing new technologies in precision medicine and making possible clinical trials that involve biologists and bioinformaticians. Mayo Clinic's Center for Individualized Medicine has set up a genomic tumor board that assesses test results from clinical genomic panel testing on patients, with further validation using array comparative genomic hybridization and clinical WES. The board recommends treatment based on the test results [280]. Institute Curie has established an information system for data integration that analyzes genomic variations in patient samples in real time and provide the test results to the clinicians and biologists to help them decide on a precision therapy. This pipeline was assessed in the context of SHIVA clinical trial that evaluates the response from targeted therapy using tumor molecular profiling vis-à-vis the conventional therapy for refractory cancer [281]. In addition to cataloging and classifying the single-nucleotide variants (SNVs), another important aspect in precision medicine is to link the SNVs to drug response. Choudhury *et al.* cataloged 2640 rarely occurring SNPs, most not yet functionally characterized. The putative deleterious SNVs identified in this study may potentially modulate the drug response even if these variations do not significantly alter the protein structure [282].

Examples of success stories of translational bioinformatics

Efforts to translate discoveries from bench to bedside have begun to yield fruits, thanks largely to the advances in translational bioinformatics. Tenenbaum (2016) presents some success stories of translational bioinformatics in precision medicine. Next-generation sequencing and bioinformatics unraveled a mutation and a novel copy number variation in the TTN gene in a newborn suffering from long QT syndrome (LQTS), a condition impacting heartbeat, at the Stanford's Lucile Packard Children's Hospital. A 14-year-old boy with severe combined immunodeficiency was eventually diagnosed for leptospirosis after his spinal fluid was subjected to sequencing and bioinformatics that revealed hundreds of reads matching the *Leptospira* sequences. Another success story is the FoundationOne test that examines the tumors for known mutations via next-generation sequencing. This test made possible targeted treatment for ~20 cancer patients [283]. Perhaps, because of the promise this new technology holds in the early detection of cancer, significant efforts have gone into deciphering genomic alternations or other biomarkers in different types of cancer, such as mRNA/miRNA, cytokine or urine-based metabolites in lung cancer [195], or novel mutations, fusions or copy number variations in biliary tract cancer [284]. A recently developed evidence-based software TREATMENTMAP used a panel of pharmacogenomic markers to probe the genomic sequences from pancreatic tumors; this diagnostic tool was shown to identify known driver mutations as well as biomarkers for effective treatment [285].

Conclusions and future directions

Metabolomics is a powerful 'omics' approach that has already found its sturdy place in medical applications. It has an intrinsic power to become one of the major components of precision medicine, and the number of metabolomics applications is rapidly growing. We believe that to fully unleash the power of metabolomics, several challenges in technology, informatics and logistics have to be addressed. Because multiple technologies and a large array of different instruments are used for data generation, standardization is becoming increasingly important. Standard procedures for sample collection and biobanking, data formats, data analysis and shared repositories have to be adopted by both research and medical communities. In the near future, we expect to see significant improvement in imaging technologies combined with prediction algorithms to be able to differentiate between healthy and diseased tissue or identify heterogeneity within diseased organ or tissue. Medical databases incorporating metabolomics data have to be developed and populated with large-scale metabotyping data for various diseases. In the bioinformatics field, we foresee major improvement in real time analysis of different layers of omics data sampled from patients and broad incorporation of machine learning and AI systems to provide physicians with fully automated clinical analyzers capable in assisting in disease diagnosis, devising treatment strategy and predicting prognosis.

Key Points

- Precision medicine relies heavily on developments in systems biology and omics disciplines, including metabolomics.
- The combination of analytical techniques is required for precision medicine metabotyping.
- Metabolomics for precision medicine require specialized computational and big data management tools.
- Future developments in metabolomics for precision medicine will require broad incorporation of machine learning techniques and AI systems to provide physicians with fully automated clinical analyzers capable in assisting in disease diagnosis, devising treatment strategy and predicting prognosis.

Funding

References

1. Shulaev V. Metabolomics technology and bioinformatics. *Brief Bioinform* 2006;**7**(2):128–39.
2. Halket JM, Waterman D, Przyborowska AM. Chemical derivatization and mass spectral libraries in metabolic profiling by GC/MS and LC/MS/MS. *J Exp Bot* 2005;**56**(410):219–43.
3. Spratlin JL, Serkova NJ, Eckhardt SG. Clinical applications of metabolomics in oncology: a review. *Clin Cancer Res* 2009;**15**(2):431–40.
4. Kennedy AD, Pappan KL, Donti TR, *et al.* Elucidation of the complex metabolic profile of cerebrospinal fluid using an untargeted biochemical profiling assay. *Mol Genet Metab* 2017;**121**(2):83–90.

5. Hampe MH, Panaskar SN, Yadav AA, *et al*. Gas chromatography/mass spectrometry-based urine metabolome study in children for inborn errors of metabolism: an Indian experience. *Clin Biochem* 2017;**50**(3):121–6.

6. Tebani A, Abily-Donval L, Afonso C, *et al*. Clinical metabolomics: the new metabolic window for inborn errors of metabolism investigations in the post-genomic era. *Int J Mol Sci* 2016;**17**(7):1167.

7. Jelliffe-Pawlowski LL, Norton ME, Baer RJ, *et al*. Gestational dating by metabolic profile at birth: a California cohort study. *Am J Obstet Gynecol* 2016;**214**(4):511.e1–e13.

8. Vernon HJ. Inborn errors of metabolism: advances in diagnosis and therapy. *JAMA Pediatr* 2015;**169**(8):778–82.

9. Creek DJ, Nijagal B, Kim DH, *et al*. Metabolomics guides rational development of a simplified cell culture medium for drug screening against Trypanosoma brucei. *Antimicrob Agents Chemother* 2013;**57**(6):2768–79.

10. Harrigan GG, Yates LA. High-throughput screening, metabolomics and drug discovery. *IDrugs* 2006;**9**(3):188–92.

11. Lin YS, Kerr SJ, Randolph T, *et al*. Prediction of intravenous busulfan clearance by endogenous plasma biomarkers using global pharmacometabolomics. *Metabolomics* 2016;**12**(10):161.

12. Katsila T, Matsoukas MT, Patrinos GP, *et al*. Pharmacometabolomics informs quantitative radiomics for glioblastoma diagnostic innovation. *OMICS* 2017;**21**:429–39.

13. Kim B, Lee JW, Hong KT, *et al*. Pharmacometabolomics for predicting variable busulfan exposure in paediatric haematopoietic stem cell transplantation patients. *Sci Rep* 2017;**7**(1):1711.

14. Amin AM, Sheau Chin L, Azri Mohamed Noor D, *et al*. The personalization of clopidogrel antiplatelet therapy: the role of integrative pharmacogenetics and pharmacometabolomics. *Cardiol Res Pract* 2017;**2017**:8062796.

15. Kantae V, Krekels EHJ, Esdonk MJV, *et al*. Integration of pharmacometabolomics with pharmacokinetics and pharmacodynamics: towards personalized drug therapy. *Metabolomics* 2017;**13**(1):9.

16. Balasopoulou A, Patrinos GP, Katsila T. Pharmacometabolomics informs viromics toward precision medicine. *Front Pharmacol* 2016;**7**:411.

17. Au A, Cheng KK, Wei LK. Metabolomics, lipidomics and pharmacometabolomics of human hypertension. *Adv Exp Med Biol* 2017;**956**:599–613.

18. Neavin D, Kaddurah-Daouk R, Weinshilboum R. Pharmacometabolomics informs pharmacogenomics. *Metabolomics* 2016;**12**:121.

19. Hornischer K, Häußler S. Diagnostics and resistance profiling of bacterial pathogens. *Curr Top Microbiol Immunol* 2016;**398**:89–102.

20. Gilbert JA, Quinn RA, Debelius J, *et al*. Microbiome-wide association studies link dynamic microbial consortia to disease. *Nature* 2016;**535**(7610):94–103.

21. Mussap M. Laboratory medicine in neonatal sepsis and inflammation. *J Matern Fetal Neonatal Med* 2012;**25(Suppl 4)**:32–4.

22. Niklinski J, Kretowski A, Moniuszko M, *et al*. Systematic biobanking, novel imaging techniques, and advanced molecular analysis for precise tumor diagnosis and therapy: the Polish MOBIT project. *Adv Med Sci* 2017;**62**(2):405–13.

23. Pandey R, Caflisch L, Lodi A, *et al*. Metabolomic signature of brain cancer. *Mol Carcinog* 2017;**56**(11):2355–71.

24. Verma V, Simone CB, II, Krishnan S, *et al*. The rise of radiomics and implications for oncologic management. *J Natl Cancer Inst* 2017;**109**(7). doi:10.1093/jnci/djx055.

25. Ghasemi M, Nabipour I, Omrani A, *et al*. Precision medicine and molecular imaging: new targeted approaches toward cancer therapeutic and diagnosis. *Am J Nucl Med Mol Imaging* 2016;**6**:310–27.

26. Jenior ML, Leslie JL, Young VB, *et al*. *Clostridium difficile* colonizes alternative nutrient niches during infection across distinct murine gut microbiomes. *mSystems* 2017;**2**:e00063–17.

27. Blum HE. The human microbiome. *Adv Med Sci* 2017;**62**(2):414–20.

28. Daliri EB, Wei S, Oh DH, *et al*. The human microbiome and metabolomics: current concepts and applications. *Crit Rev Food Sci Nutr* 2017;**57**(16):3565–76.

29. Zhang C, Zhao L. Strain-level dissection of the contribution of the gut microbiome to human metabolic disease. *Genome Med* 2016;**8**(1):41.

30. Clish CB. Metabolomics: an emerging but powerful tool for precision medicine. *Cold Spring Harb Mol Case Stud* 2015;**1**(1):a000588.

31. Beger RD, Dunn W, Schmidt MA, *et al*. Metabolomics enables precision medicine: "a white paper, community perspective". *Metabolomics* 2016;**12**:149.

32. Eckerle M, Ambroggio L, Puskarich M, *et al*. Metabolomics as a driver in advancing precision medicine in sepsis. *Pharmacotherapy* 2017;**37**(9):1023–32.

33. Trivedi DK, Hollywood KA, Goodacre R. Metabolomics for the masses: the future of metabolomics in a personalized world. *New Horiz Transl Med* 2017;**3**(6):294–305.

34. Lam SM, Wang Y, Li B, *et al*. Metabolomics through the lens of precision cardiovascular medicine. *J Genet Genomics* 2017;**44**(3):127–38.

35. Wishart DS. Emerging applications of metabolomics in drug discovery and precision medicine. *Nat Rev Drug Discov* 2016;**15**(7):473–84.

36. Gillman MW, Hammond RA. Precision treatment and precision prevention: integrating "below and above the skin". *JAMA Pediatr* 2016;**170**(1):9–10.

37. Thomas DC. What does "Precision Medicine" have to say about prevention? *Epidemiology* 2017;**28**(4):479–83.

38. Semmar N. Metabotype concept: flexibility, usefulness and meaning in different biological populations. In: U. Roessner (ed) *Metabolomics*. InTech, 2012, 131–66.

39. Brennan L. Use of metabotyping for optimal nutrition. *Curr Opin Biotechnol* 2017;**44**:35–8.

40. Holmes E, Wilson ID, Nicholson JK. Metabolic phenotyping in health and disease. *Cell* 2008;**134**(5):714–17.

41. Esterhuizen K, van der Westhuizen FH, Louw R. Metabolomics of mitochondrial disease. *Mitochondrion* 2017;**35**:97–110.

42. Hocher B, Adamski J. Metabolomics for clinical use and research in chronic kidney disease. *Nat Rev Nephrol* 2017;**13**(5):269–84.

43. Dona AC, Coffey S, Figtree G. Translational and emerging clinical applications of metabolomics in cardiovascular disease diagnosis and treatment. *Eur J Prev Cardiol* 2016;**23**(15):1578–89.

44. James EL, Parkinson EK. Serum metabolomics in animal models and human disease. *Curr Opin Clin Nutr Metab Care* 2015;**18**(5):478–83.

45. Rhee EP. Metabolomics and renal disease. *Curr Opin Nephrol Hypertens* 2015;**24**(4):371–9.

46. Shariff MIF, Kim JU, Ladep NG, *et al*. The plasma and serum metabotyping of hepatocellular carcinoma in a Nigerian and Egyptian cohort using proton nuclear magnetic resonance spectroscopy. *J Clin Exp Hepatol* 2017;**7**(2):83–92.

47. Ladep NG, Dona AC, Lewis MR, *et al*. Discovery and validation of urinary metabotypes for the diagnosis of hepatocellular carcinoma in West Africans. *Hepatology* 2014;**60**(4): 1291–301.

48. Jacobs JP, Goudarzi M, Singh N, *et al*. A disease-associated microbial and metabolomics state in relatives of pediatric Inflammatory Bowel Disease patients. *Cell Mol Gastroenterol Hepatol* 2016;**2**(6):750–66.

49. Martin FP, Ezri J, Cominetti O, *et al*. Urinary metabolic phenotyping reveals differences in the metabolic status of healthy and Inflammatory Bowel Disease (IBD) children in relation to growth and disease activity. *Int J Mol Sci* 2016;**17**(8):1310.

50. Reinke SN, Gallart-Ayala H, Gomez C, *et al*. Metabolomics analysis identifies different metabotypes of asthma severity. *Eur Respir J* 2017;**49**(3):1601740.

51. Dumas ME, Domange C, Calderari S, *et al*. Topological analysis of metabolic networks integrating co-segregating transcriptomes and metabolomes in type 2 diabetic rat congenic series. *Genome Med* 2016;**8**(1):101.

52. Wolahan SM, Hirt D, Glenn TC. Frontiers in neuroengineering: translational metabolomics of head injury: exploring dysfunctional cerebral metabolism with *ex vivo* NMR spectroscopy-based metabolite quantification. In: F. H. Kobeissy (ed) *Brain Neurotrauma: Molecular, Neuropsychological, and Rehabilitation Aspects*. Boca Raton, FL: CRC Press/Taylor & Francis (c); LLC, 2015.

53. Dumas ME, Kinross J, Nicholson JK. Metabolic phenotyping and systems biology approaches to understanding metabolic syndrome and fatty liver disease. *Gastroenterology* 2014;**146**(1):46–62.

54. Luan H, Liu LF, Tang Z, *et al*. Comprehensive urinary metabolomic profiling and identification of potential noninvasive marker for idiopathic Parkinson's disease. *Sci Rep* 2015;**5**(1): 13888.

55. Calvani R, Miccheli A, Capuani G, *et al*. Gut microbiome-derived metabolites characterize a peculiar obese urinary metabotype. *Int J Obes* 2010;**34**(6):1095–8.

56. Tam ZY, Ng SP, Tan LQ, *et al*. Metabolite profiling in identifying metabolic biomarkers in older people with late-onset type 2 diabetes mellitus. *Sci Rep* 2017;**7**:4392.

57. Hu J, Locasale JW, Bielas JH, *et al*. Heterogeneity of tumor-induced gene expression changes in the human metabolic network. *Nat Biotechnol* 2013;**31**(6):522–9.

58. Hensley CT, Faubert B, Yuan Q, *et al*. Metabolic heterogeneity in human lung tumors. *Cell* 2016;**164**(4):681–94.

59. Colgan SP, Curtis VF, Campbell EL. The inflammatory tissue microenvironment in IBD. *Inflamm Bowel Dis* 2013;**19**(10): 2238–44.

60. Byun JS, Yi HS. Hepatic immune microenvironment in alcoholic and nonalcoholic liver disease. *Biomed Res Int* 2017; **2017**:6862439.

61. Viil J, Maasalu K, Maemets-Allas K, *et al*. Laminin-rich blood vessels display activated growth factor signaling and act as the proliferation centers in Dupuytren's contracture. *Arthritis Res Ther* 2015;**17**(1):144.

62. Hedigan K. Fibrotic disease: targeting the microenvironment. *Nat Rev Drug Discov* 2010;**9**:840–1.

63. Rahat MA, Shakya J. Parallel aspects of the microenvironment in cancer and autoimmune disease. *Mediators Inflamm* 2016;**2016**:4375120.

64. Fuster JJ, Ouchi N, Gokce N, *et al*. Obesity-induced changes in adipose tissue microenvironment and their impact on cardiovascular disease. *Circ Res* 2016;**118**(11):1786–807.

65. Okegawa T, Morimoto M, Nishizawa S, *et al*. Intratumor heterogeneity in primary kidney cancer revealed by metabolic profiling of multiple spatially separated samples within tumors. *EBioMedicine* 2017;**19**:31–8.

66. Yap TA, Gerlinger M, Futreal PA, *et al*. Intratumor heterogeneity: seeing the wood for the trees. *Sci Transl Med* 2012; **4**(127):127ps10.

67. Gerlinger M, Rowan AJ, Horswell S, *et al*. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med* 2012;**366**:883–92.

68. Zenobi R. Single-cell metabolomics: analytical and biological perspectives. *Science* 2013;**342**(6163):1243259.

69. Liu X, Locasale JW. Metabolomics reveals intratumor heterogeneity—implications for precision medicine. *EBioMedicine* 2017;**19**:4–5.

70. Larsen PE, Dai Y. Metabolome of human gut microbiome is predictive of host dysbiosis. *Gigascience* 2015;**4**:42.

71. Stiegel MA, Pleil JD, Sobus JR, *et al*. Linking physiological parameters to perturbations in the human exposome: environmental exposures modify blood pressure and lung function via inflammatory cytokine pathway. *J Toxicol Environ Health A* 2017;**80**(9):485–501.

72. Niedzwiecki MM, Miller GW. The exposome paradigm in human health: lessons from the Emory exposome summer course. *Environ Health Perspect* 2017;**125**(6):064502.

73. Buck Louis GM, Smarr MM, Patel CJ. The exposome research paradigm: an opportunity to understand the environmental basis for human health and disease. *Curr Environ Health Rep* 2017;**4**(1):89–98.

74. Andra SS, Austin C, Patel D, *et al*. Trends in the application of high-resolution mass spectrometry for human biomonitoring: an analytical primer to studying the environmental chemical space of the human exposome. *Environ Int* 2017; **100**:32–61.

75. Athersuch TJ, Keun HC. Metabolic profiling in human exposome studies. *Mutagenesis* 2015;**30**(6):755–62.

76. Vrijheid M, Slama R, Robinson O, *et al*. The human early-life exposome (HELIX): project rationale and design. *Environ Health Perspect* 2014;**122**:535–44.

77. Walsh CJ, Guinane CM, Hill C, *et al*. In silico identification of bacteriocin gene clusters in the gastrointestinal tract, based on the Human Microbiome Project's reference genome database. *BMC Microbiol* 2015;**15**(1):183.

78. Chen T, Yu WH, Izard J, *et al*. The human oral microbiome database: a web accessible resource for investigating oral microbe taxonomic and genomic information. *Database* 2010;**2010**(0):baq013.

79. Wishart DS, Mandal R, Stanislaus A, *et al*. Cancer metabolomics and the human metabolome database. *Metabolites* 2016;**6**(1):10.

80. Wishart DS, Jewison T, Guo AC, *et al*. HMDB 3.0–the human metabolome database in 2013. *Nucleic Acids Res* 2013;**41**: D801–7.

81. Forsythe IJ, Wishart DS. Exploring human metabolites using the human metabolome database. *Curr Protoc Bioinformatics* 2009;**Chapter 14**:Unit14.8.

82. Wishart DS. Human metabolome database: completing the 'human parts list'. *Pharmacogenomics* 2007;**8**(7):683–6.

83. Wishart DS, Tzur D, Knox C, *et al*. HMDB: the human metabolome database. *Nucleic Acids Res* 2007;**35**:D521–6.

84. Zheng H, Ji J, Zhao L, *et al*. Prediction and diagnosis of renal cell carcinoma using nuclear magnetic resonance-based serum metabolomics and self-organizing maps. *Oncotarget* 2016;**7**(37):59189–98.

85. Wang D, Wang X, Kong J, *et al*. GC-MS-based metabolomics discovers a shared serum metabolic characteristic among three types of epileptic seizures. *Epilepsy Res* 2016;**126**:83–9.

86. Chen J, Hou W, Han B, *et al*. Target-based metabolomics for the quantitative measurement of 37 pathway metabolites in rat brain and serum using hydrophilic interaction ultra-high-performance liquid chromatography-tandem mass spectrometry. *Anal Bioanal Chem* 2016;**408**(10):2527–42.

87. West PR, Amaral DG, Bais P, *et al*. Metabolomics as a tool for discovery of biomarkers of autism spectrum disorder in the blood plasma of children. *PLoS One* 2014;**9**(11):e112445.

88. Bala L, Sharma A, Yellapa RK, *et al*. (1)H NMR spectroscopy of ascitic fluid: discrimination between malignant and benign ascites and comparison of the results with conventional methods. *NMR Biomed* 2008;**21**(6):606–14.

89. Yilmaz A, Geddes T, Han B, *et al*. Diagnostic biomarkers of Alzheimer's disease as identified in saliva using 1H NMR-based metabolomics. *J Alzheimers Dis* 2017;**58**(2):355–9.

90. Figueira J, Gouveia-Figueira S, Ohman C, *et al*. Metabolite quantification by NMR and LC-MS/MS reveals differences between unstimulated, stimulated, and pure parotid saliva. *J Pharm Biomed Anal* 2017;**140**:295–300.

91. Figueira J, Jonsson P, Nordin Adolfsson A, *et al*. NMR analysis of the human saliva metabolome distinguishes dementia patients from matched controls. *Mol Biosyst* 2016;**12**(8):2562–71.

92. Chen L, Zhou L, Chan EC, *et al*. Characterization of the human tear metabolome by LC-MS/MS. *J Proteome Res* 2011;**10**(10):4876–82.

93. Surowiec I, Karimpour M, Gouveia-Figueira S, *et al*. Multi-platform metabolomics assays for human lung lavage fluids in an air pollution exposure study. *Anal Bioanal Chem* 2016;**408**(17):4751–64.

94. Lynch MJ, Masters J, Pryor JP, *et al*. Ultra high field NMR spectroscopic studies on human seminal fluid, seminal vesicle and prostatic secretions. *J Pharm Biomed Anal* 1994;**12**(1):5–19.

95. Stoll ML, Kumar R, Lefkowitz EJ, *et al*. Fecal metabolomics in pediatric spondyloarthritis implicate decreased metabolic diversity and altered tryptophan metabolism as pathogenic factors. *Genes Immun* 2016;**17**(7):400–5.

96. Winter T, Lange A, Hannemann A, *et al*. Contamination of dried blood spots—an underestimated risk in newborn screening. *Clin Chem Lab Med* 2017. doi:10.1515/cclm-2017-0270.

97. Imai A, Kishita Y, Nakayama Y, *et al*. Dried blood spots for newborn screening allows easy determination of a high heteroplasmy rate in severe infantile cardiomyopathy. *Int J Cardiol* 2016;**221**:446–9.

98. Jung S, Tran NT, Gospe SM, Jr, *et al*. Preliminary investigation of the use of newborn dried blood spots for screening pyridoxine-dependent epilepsy by LC-MS/MS. *Mol Genet Metab* 2013;**110**:237–40.

99. Shokry E, Villanelli F, Malvagia S, *et al*. Therapeutic drug monitoring of carbamazepine and its metabolite in children from dried blood spots using liquid chromatography and tandem mass spectrometry. *J Pharm Biomed Anal* 2015;**109**:164–70.

100. Hofman S, Bolhuis MS, Koster RA, *et al*. Role of therapeutic drug monitoring in pulmonary infections: use and potential for expanded use of dried blood spot samples. *Bioanalysis* 2015;**7**(4):481–95.

101. Tran C, Yazdanpanah M, Kyriakopoulou L, *et al*. Stable isotope dilution microquantification of creatine metabolites in plasma, whole blood and dried blood spots for pharmacological studies in mouse models of creatine deficiency. *Clin Chim Acta* 2014;**436**:160–8.

102. Ter Heine R, Mulder JW, van Gorp EC, *et al*. Clinical evaluation of the determination of plasma concentrations of darunavir, etravirine, raltegravir and ritonavir in dried blood spot samples. *Bioanalysis* 2011;**3**(10):1093–7.

103. Drolet J, Tolstikov V, Williams BA, *et al*. Integrated metabolomics assessment of human dried blood spots and urine strips. *Metabolites* 2017;**7**:pii:E35.

104. de Sain-van der Velden MGM, van der Ham M, Gerrits J, *et al*. Quantification of metabolites in dried blood spots by direct infusion high resolution mass spectrometry. *Anal Chim Acta* 2017;**979**:45–50.

105. Hernandes VV, Barbas C, Dudzik D. A review of blood sample handling and pre-processing for metabolomics studies. *Electrophoresis* 2017;**38**(18):2232–41.

106. Patejko M, Jacyna J, Markuszewski MJ. Sample preparation procedures utilized in microbial metabolomics: an overview. *J Chromatogr B Analyt Technol Biomed Life Sci* 2017;**1043**:150–7.

107. Wu Y, Li L. Sample normalization methods in quantitative metabolomics. *J Chromatogr A* 2016;**1430**:80–95.

108. Mastrangelo A, Ferrarini A, Rey-Stolle F, *et al*. From sample treatment to biomarker discovery: a tutorial for untargeted metabolomics based on GC-(EI)-Q-MS. *Anal Chim Acta* 2015;**900**:21–35.

109. Bradburne C, Graham D, Kingston HM, *et al*. Overview of 'Omics' technologies for military occupational health surveillance and medicine. *Mil Med* 2015;**180(10 Suppl)**:34–48.

110. Abuja PM, Ehrhart F, Schoen U, *et al*. Alterations in human liver metabolome during prolonged cryostorage. *J Proteome Res* 2015;**14**(7):2758–68.

111. LaBaer J. Improving international research with clinical specimens: 5 achievable objectives. *J Proteome Res* 2012;**11**(12):5592–601.

112. Yang L, Chen Y, Yu C, *et al*. Biobanks and their clinical application and informatics challenges. *Adv Exp Med Biol* 2016;**939**:241–57.

113. Tan SY, Sandanaraj E, Tang C, *et al*. Biobanking: an important resource for precision medicine in glioblastoma. *Adv Exp Med Biol* 2016;**951**:47–56.

114. Felmeister AS, Masino AJ, Rivera TJ, *et al*. The biorepository portal toolkit: an honest brokered, modular service oriented software tool set for biospecimen-driven translational research. *BMC Genomics* 2016;**17**(S4):434.

115. Fiehn O. Metabolomics–the link between genotypes and phenotypes. *Plant Mol Biol* 2002;**48**(1–2):155–71.

116. Fiehn O. Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks. *Comp Funct Genomics* 2001;**2**(3):155–68.

117. Sumner LW, Amberg A, Barrett D, *et al*. Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics* 2007;**3**(3):211–21.

118. Far J, Delvaux C, Kune C, *et al*. The use of ion mobility mass spectrometry for isomer composition determination extracted from Se-rich yeast. *Anal Chem* 2014;**86**(22):11246–54.

119. Park VY, Yoon D, Koo JS, *et al*. Intratumoral agreement of high-resolution magic angle spinning magnetic resonance spectroscopic profiles in the metabolic characterization of breast cancer. *Medicine* 2016;**95**(15):e3398.

120. Precht C, Diserens G, Oevermann A, *et al*. Visibility of lipid resonances in HR-MAS spectra of brain biopsies subject to

spinning rate variation. *Biochim Biophys Acta* 2015;**1851**(12): 1539–44.

121. Martinez-Bisbal MC, Martinez-Granados B, Rovira V, *et al*. Magnetic resonance spectroscopy and imaging on fresh human brain tumor biopsies at microscopic resolution. *Anal Bioanal Chem* 2015;**407**:6771–80.

122. Monleón D, Morales JM, Gonzalez-Darder J, *et al*. Benign and atypical meningioma metabolic signatures by high-resolution magic-angle spinning molecular profiling. *J Proteome Res* 2008; **7**(7):2882–8.

123. Chan EC, Koh PK, Mal M, *et al*. Metabolic profiling of human colorectal cancer using high-resolution magic angle spinning nuclear magnetic resonance (HR-MAS NMR) spectroscopy and gas chromatography mass spectrometry (GC/MS). *J Proteome Res* 2009;**8**(1):352–61.

124. Palmnas MS, Vogel HJ. The future of NMR metabolomics in cancer therapy: towards personalizing treatment and developing targeted drugs? *Metabolites* 2013;**3**(2):373–96.

125. Weiss RH, Kim K. Metabolomics in the study of kidney diseases. *Nat Rev Nephrol* 2011;**8**(1):22–33.

126. Dessi A, Corsello G, Stronati M, *et al*. New diagnostic possibilities in systemic neonatal infections: metabolomics. *Early Hum Dev* 2014;**90(Suppl 1)**:S19–21.

127. Soininen P, Kangas AJ, Wurtz P, *et al*. Quantitative serum nuclear magnetic resonance metabolomics in cardiovascular epidemiology and genetics. *Circ Cardiovasc Genet* 2015; **8**(1):192–206.

128. Rankin NJ, Preiss D, Welsh P, *et al*. The emergence of proton nuclear magnetic resonance metabolomics in the cardiovascular arena as viewed from a clinical perspective. *Atherosclerosis* 2014;**237**(1):287–300.

129. Rhee EP, Gerszten RE. Metabolomics and cardiovascular biomarker discovery. *Clin Chem* 2012;**58**(1):139–47.

130. Stringer KA, McKay RT, Karnovsky A, *et al*. Metabolomics and its application to acute lung diseases. *Front Immunol* 2016;**7**:44.

131. Fanos V, Barberini L, Antonucci R, *et al*. Pharma-metabolomics in neonatology: is it a dream or a fact? *Curr Pharm Des* 2012;**18**(21):2996–3006.

132. Perez-Rambla C, Puchades-Carrasco L, Garcia-Flores M, *et al*. Non-invasive urinary metabolomic profiling discriminates prostate cancer from benign prostatic hyperplasia. *Metabolomics* 2017;**13**:52.

133. Watanabe M, Sheriff S, Lewis KB, *et al*. Metabolic profiling comparison of human pancreatic ductal epithelial cells and three pancreatic cancer cell lines using NMR based metabonomics. *J Mol Biomark Diagn* 2012;**3**(2):S3-002.

134. Wang L, Chen J, Chen L, *et al*. 1H-NMR based metabonomic profiling of human esophageal cancer tissue. *Mol Cancer* 2013;**12**:25.

135. Davis VW, Schiller DE, Eurich D, *et al*. Urinary metabolomic signature of esophageal cancer and Barrett's esophagus. *World J Surg Oncol* 2012;**10**:271.

136. Zhang X, Xu L, Shen J, *et al*. Metabolic signatures of esophageal cancer: NMR-based metabolomics and UHPLC-based focused metabolomics of blood serum. *Biochim Biophys Acta* 2013;**1832**(8):1207–16.

137. Singh A, Sharma RK, Chagtoo M, *et al*. 1H NMR metabolomics reveals association of high expression of inositol 1, 4, 5 trisphosphate receptor and metabolites in breast cancer patients. *PLoS One* 2017;**12**(1):e0169330.

138. Wang H, Zhang H, Deng P, *et al*. Tissue metabolic profiling of human gastric cancer assessed by (1)H NMR. *BMC Cancer* 2016;**16**(1):371.

139. Wang H, Wang L, Zhang H, *et al*. (1)H NMR-based metabolic profiling of human rectal cancer tissue. *Mol Cancer* 2013; **12**(1):121.

140. Rocha CM, Barros AS, Gil AM, *et al*. Metabolic profiling of human lung cancer tissue by 1H high resolution magic angle spinning (HRMAS) NMR spectroscopy. *J Proteome Res* 2010; **9**(1):319–32.

141. Zhang X, Zhu X, Wang C, *et al*. Non-targeted and targeted metabolomics approaches to diagnosing lung cancer and predicting patient prognosis. *Oncotarget* 2016;**7**(39):63437–48.

142. Engskog M, Bjorklund M, Haglof J, *et al*. Metabolic profiling of epithelial ovarian cancer cell lines: evaluation of harvesting protocols for profiling using NMR spectroscopy. *Bioanalysis* 2015;**7**(2):157–66.

143. Monteiro MS, Barros AS, Pinto J, *et al*. Nuclear Magnetic Resonance metabolomics reveals an excretory metabolic signature of renal cell carcinoma. *Sci Rep* 2016;**6**(1):37275.

144. Zhang H, Qiao L, Li X, *et al*. Tissue metabolic profiling of lymph node metastasis of colorectal cancer assessed by 1H NMR. *Oncol Rep* 2016;**36**(6):3436–48.

145. Junot C, Madalinski G, Tabet JC, *et al*. Fourier transform mass spectrometry for metabolome analysis. *Analyst* 2010;**135**(9): 2203–19.

146. Junot C, Fenaille F, Colsch B, *et al*. High resolution mass spectrometry based techniques at the crossroads of metabolic pathways. *Mass Spectrom Rev* 2014;**33**(6):471–500.

147. Wei H, Pasman W, Rubingh C, *et al*. Urine metabolomics combined with the personalized diagnosis guided by Chinese medicine reveals subtypes of pre-diabetes. *Mol Biosyst* 2012;**8**(5):1482–91.

148. Liu S, Wang L, Hu C, *et al*. Plasma metabolomics profiling of maintenance hemodialysis based on capillary electrophoresis - time of flight mass spectrometry. *Sci Rep* 2017;**7**(1):8150.

149. Prinsen HC, Schiebergen-Bronkhorst BG, Roeleveld MW, *et al*. Rapid quantification of underivatized amino acids in plasma by hydrophilic interaction liquid chromatography (HILIC) coupled with tandem mass-spectrometry. *J Inherit Metab Dis* 2016;**39**(5):651–60.

150. Domingues DS, Crevelin EJ, de Moraes LA, *et al*. Simultaneous determination of amino acids and neurotransmitters in plasma samples from schizophrenic patients by hydrophilic interaction liquid chromatography with tandem mass spectrometry. *J Sep Sci* 2015;**38**(5):780–7.

151. Methlie P, Hustad S, Kellman R, *et al*. Multisteroid LC–MS/MS assay for glucocorticoids and androgens and its application in Addison's disease. *Endocr Connect* 2013;**2**(3):125–36.

152. Domingues DS, Souza ID, Queiroz ME. Analysis of drugs in plasma samples from schizophrenic patients by column-switching liquid chromatography-tandem mass spectrometry with organic-inorganic hybrid cyanopropyl monolithic column. *J Chromatogr B Analyt Technol Biomed Life Sci* 2015; **993–4**:26–35.

153. Ponnuru VS, Challa BR, Nadendla R. Quantification of sibutramine and its two metabolites in human plasma by LC–ESI-MS/MS and its application in a bioequivalence study. *J Pharm Anal* 2012;**2**(4):249–57.

154. Ghaste M, Mistrik R, Shulaev V. Applications of Fourier Transform Ion Cyclotron Resonance (FT-ICR) and orbitrap based high resolution mass spectrometry in metabolomics and lipidomics. *Int J Mol Sci* 2016;**17**(6):17.

155. Brown SC, Kruppa G, Dasseux JL. Metabolomics applications of FT-ICR mass spectrometry. *Mass Spectrom Rev* 2005;**24**(2): 223–31.

156. Gray N, Lewis MR, Plumb RS, *et al.* High-throughput microbore UPLC-MS metabolic phenotyping of urine for large-scale epidemiology studies. *J Proteome Res* 2015;**14**(6): 2714–21.

157. Ni Q, Reid KR, Burant CF, *et al.* Capillary LC-MS for high sensitivity metabolomic analysis of single islets of Langerhans. *Anal Chem* 2008;**80**(10):3539–46.

158. Edwards JL, Edwards RL, Reid KR, *et al.* Effect of decreasing column inner diameter and use of off-line two-dimensional chromatography on metabolite detection in complex mixtures. *J Chromatogr A* 2007;**1172**(2):127–34.

159. Jackson SN, Baldwin K, Muller L, *et al.* Imaging of lipids in rat heart by MALDI-MS with silver nanoparticles. *Anal Bioanal Chem* 2014;**406**(5):1377–86.

160. Angel PM, Bayoumi AS, Hinton RB, *et al.* MALDI imaging mass spectrometry as a lipidomic approach to heart valve research. *J Heart Valve Dis* 2016;**25**:240–52.

161. Angel PM, Baldwin HS, Gottlieb Sen D, *et al.* Advances in MALDI imaging mass spectrometry of proteins in cardiac tissue, including the heart valve. *Biochim Biophys Acta* 2017; **1865**(7):927–35.

162. Mohammadi AS, Phan NT, Fletcher JS, *et al.* Intact lipid imaging of mouse brain samples: MALDI, nanoparticle-laser desorption ionization, and 40 keV argon cluster secondary ion mass spectrometry. *Anal Bioanal Chem* 2016;**408**(24): 6857–68.

163. Ronci M, Sharma S, Chataway T, *et al.* MALDI-MS-imaging of whole human lens capsule. *J Proteome Res* 2011;**10**(8):3522–9.

164. Jiao J, Miao A, Zhang Y, *et al.* Imaging phosphorylated peptide distribution in human lens by MALDI MS. *Analyst* 2015; **140**(12):4284–90.

165. Bhandari DR, Schott M, Rompp A, *et al.* Metabolite localization by atmospheric pressure high-resolution scanning microprobe matrix-assisted laser desorption/ionization mass spectrometry imaging in whole-body sections and individual organs of the rove beetle Paederus riparius. *Anal Bioanal Chem* 2015;**407**(8):2189–201.

166. Rao T, Shao Y, Hamada N, *et al.* Pharmacokinetic study based on a matrix-assisted laser desorption/ionization quadrupole ion trap time-of-flight imaging mass microscope combined with a novel relative exposure approach: a case of octreotide in mouse target tissues. *Anal Chim Acta* 2017;**952**:71–80.

167. Zhang J, Feider CL, Nagi C, *et al.* Detection of metastatic breast and thyroid cancer in lymph nodes by desorption electrospray ionization mass spectrometry imaging. *J Am Soc Mass Spectrom* 2017;**28**(6):1166–74.

168. Banerjee S, Zare RN, Tibshirani RJ, *et al.* Diagnosis of prostate cancer by desorption electrospray ionization mass spectrometric imaging of small metabolites and lipids. *Proc Natl Acad Sci USA* 2017;**114**(13):3334–9.

169. Inglese P, McKenzie JS, Mroz A, *et al.* Deep learning and 3D-DESI imaging reveal the hidden metabolic heterogeneity of cancer. *Chem Sci* 2017;**8**(5):3500–11.

170. Tillner J, Wu V, Jones EA, *et al.* Faster, more reproducible DESI-MS for biological tissue imaging. *J Am Soc Mass Spectrom* 2017;**28**(10):2090–8.

171. Bokhart MT, Rosen E, Thompson C, *et al.* Quantitative mass spectrometry imaging of emtricitabine in cervical tissue model using infrared matrix-assisted laser desorption electrospray ionization. *Anal Bioanal Chem* 2015;**407**(8):2073–84.

172. Sampson JS, Hawkridge AM, Muddiman DC. Construction of a versatile high precision ambient ionization source for direct analysis and imaging. *J Am Soc Mass Spectrom* 2008; **19**(10):1527–34.

173. Shobo A, Bratkowska D, Baijnath S, *et al.* Visualization of time-dependent distribution of rifampicin in rat brain using MALDI MSI and quantitative LCMS/MS. *Assay Drug Dev Technol* 2015;**13**(5):277–84.

174. Quiason CM, Shahidi-Latham SK. Imaging MALDI MS of dosed brain tissues utilizing an alternative analyte pre-extraction approach. *J Am Soc Mass Spectrom* 2015;**26**(6): 967–73.

175. Fulop A, Sammour DA, Erich K, *et al.* Molecular imaging of brain localization of liposomes in mice using MALDI mass spectrometry. *Sci Rep* 2016;**6**:33791.

176. Jones EE, Zhang W, Zhao X, *et al.* Tissue localization of glycosphingolipid accumulation in a gaucher disease mouse brain by LC-ESI-MS/MS and high-resolution MALDI imaging mass spectrometry. *SLAS Discov* 2017;**22**(10):1218–28.

177. Grey AC, Gelasco AK, Section J, *et al.* Molecular morphology of the chick heart visualized by MALDI imaging mass spectrometry. *Anat Rec* 2010;**293**(5):821–8.

178. Park ES, Lee JH, Hong JH, *et al.* Phosphatidylcholine alteration identified using MALDI imaging MS in HBV-infected mouse livers and virus-mediated regeneration defects. *PLoS One* 2014;**9**(8):e103955.

179. Kim YH, Fujimura Y, Sasaki M, *et al.* In situ label-free visualization of orally dosed strictinin within mouse kidney by MALDI-MS imaging. *J Agric Food Chem* 2014;**62**(38):9279–85.

180. Hart PJ, Francese S, Claude E, *et al.* MALDI-MS imaging of lipids in ex vivo human skin. *Anal Bioanal Chem* 2011;**401**(1): 115–25.

181. Rauser S, Marquardt C, Balluff B, *et al.* Classification of HER2 receptor status in breast cancer tissues by MALDI imaging mass spectrometry. *J Proteome Res* 2010;**9**(4):1854–63.

182. Dekker TJ, Balluff BD, Jones EA, *et al.* Multicenter matrix-assisted laser desorption/ionization mass spectrometry imaging (MALDI MSI) identifies proteomic differences in breast-cancer-associated stroma. *J Proteome Res* 2014;**13**: 4730–8.

183. Jiang L, Chughtai K, Purvine SO, *et al.* MALDI-mass spectrometric imaging revealing hypoxia-driven lipids and proteins in a breast tumor model. *Anal Chem* 2015;**87**(12):5947–56.

184. Vegvari A, Shavkunov AS, Fehniger TE, *et al.* Localization of tamoxifen in human breast cancer tumors by MALDI mass spectrometry imaging. *Clin Transl Med* 2016;**5**(1):10.

185. Alberts D, Pottier C, Smargiasso N, *et al.* MALDI imaging-guided microproteomic analyses of heterogeneous breast tumors - A pilot study. *Proteomics Clin Appl* 2017. doi: 10.1002/prca.201700062.

186. Buck A, Ly A, Balluff B, *et al.* High-resolution MALDI-FT-ICR MS imaging for the analysis of metabolites from formalin-fixed, paraffin-embedded clinical tissue samples. *J Pathol* 2015;**237**(1):123–32.

187. Jirasko R, Holcapek M, Kunes M, *et al.* Distribution study of atorvastatin and its metabolites in rat tissues using combined information from UHPLC/MS and MALDI-Orbitrap-MS imaging. *Anal Bioanal Chem* 2014;**406**:4601–10.

188. Balog J, Szaniszlo T, Schaefer KC, *et al.* Identification of biological tissues by rapid evaporative ionization mass spectrometry. *Anal Chem* 2010;**82**(17):7343–50.

189. Balog J, Sasi-Szabo L, Kinross J, *et al.* Intraoperative tissue identification using rapid evaporative ionization mass spectrometry. *Sci Transl Med* 2013;**5**(194):194ra193.

190. Golf O, Strittmatter N, Karancsi T, *et al.* Rapid evaporative ionization mass spectrometry imaging platform for direct

mapping from bulk tissue and bacterial growth media. *Anal Chem* 2015;**87**(5):2527–34.

191. Balog J, Kumar S, Alexander J, *et al*. *In vivo* endoscopic tissue identification by rapid evaporative ionization mass spectrometry (REIMS). *Angew Chem Int Ed Engl* 2015;**54**(38):11059–62.

192. Balog J, Perenyi D, Guallar-Hoyas C, *et al*. Identification of the species of origin for meat products by rapid evaporative ionization mass spectrometry. *J Agric Food Chem* 2016;**64**(23):4793–800.

193. Verplanken K, Stead S, Jandova R, *et al*. Rapid evaporative ionization mass spectrometry for high-throughput screening in food analysis: the case of boar taint. *Talanta* 2017;**169**:30–6.

194. St John ER, Balog J, McKenzie JS, *et al*. Rapid evaporative ionisation mass spectrometry of electrosurgical vapours for the identification of breast pathology: towards an intelligent knife for breast cancer surgery. *Breast Cancer Res* 2017;**19**(1):59.

195. Robles AI, Harris CC. Integration of multiple "OMIC" biomarkers: a precision medicine strategy for lung cancer. *Lung Cancer* 2017;**107**:50–8.

196. Blekherman G, Laubenbacher R, Cortes DF, *et al*. Bioinformatics tools for cancer metabolomics. *Metabolomics* 2011;**7**(3):329–43.

197. Katajamaa M, Oresic M. Data processing for mass spectrometry-based metabolomics. *J Chromatogr A* 2007;**1158**(1–2):318–28.

198. Madsen R, Lundstedt T, Trygg J. Chemometrics in metabolomics–a review in human disease diagnosis. *Anal Chim Acta* 2010;**659**(1–2):23–33.

199. Markley JL, Anderson ME, Cui Q, *et al*. New bioinformatics resources for metabolomics. *Pac Symp Biocomput* 2007;**12**:157–68.

200. Sumner LW, Mendes P, Dixon RA. Plant metabolomics: large-scale phytochemistry in the functional genomics era. *Phytochemistry* 2003;**62**(6):817–36.

201. Wilcoxen KM, Uehara T, Myint KT, *et al*. Practical metabolomics in drug discovery. *Expert Opin Drug Discov* 2010;**5**(3):249–63.

202. Duerr-Specht M, Goebel R, Holzinger A. Medicine and health care as a data problem: will computers become better medical doctors? In: A. Holzinger, C. Röcker, M. Ziefle (eds). *Smart Health: Open Problems and Future Challenges*. Cham: Springer International Publishing, 2015, 21–39.

203. Shulaev V, Cortes D, Miller G, *et al*. Metabolomics for plant stress response. *Physiol Plant* 2008;**132**(2):199–208.

204. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;**30**(15):2114–20.

205. Sugimoto M, Kawakami M, Robert M, *et al*. Bioinformatics tools for mass spectroscopy-based metabolomic data processing and analysis. *Curr Bioinform* 2012;**7**(1):96–108.

206. Sboner A, Elemento O. A primer on precision medicine informatics. *Brief Bioinform* 2016;**17**(1):145–53.

207. Toledo DM, Bowser MJ, Duffy-Hynes E, *et al*. Integrating genetics into subspecialty care: the pulmogene test—comprehensive testing for hereditary causes of lung disease. *Am J Respir Crit Care Med* 2014;**189**:A2175.

208. Li H, Durbin R. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics* 2009;**25**(14):1754–60.

209. Lee S, Hormozdiari F, Alkan C, *et al*. MoDIL: detecting small indels from clone-end sequencing with mixtures of distributions. *Nat Methods* 2009;**6**(7):473–4.

210. Chen K, Wallis JW, McLellan MD, *et al*. BreakDancer: an algorithm for high-resolution mapping of genomic structural variation. *Nat Methods* 2009;**6**(9):677–81.

211. Fromer M, Moran JL, Chambert K, *et al*. Discovery and statistical genotyping of copy-number variation from whole-exome sequencing depth. *Am J Hum Genet* 2012;**91**(4):597–607.

212. Trapnell C, Roberts A, Goff L, *et al*. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* 2012;**7**(3):562–78.

213. Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol* 2010;**11**(10):R106.

214. Cheng T, Zhan X. Pattern recognition for predictive, preventive, and personalized medicine in cancer. *EPMA J* 2017;**8**(1):51–60.

215. Ferte C, Trister AD, Huang E, *et al*. Impact of bioinformatic procedures in the development and translation of high-throughput molecular classifiers in oncology. *Clin Cancer Res* 2013;**19**(16):4315–25.

216. Lindon JC, Nicholson JK. The emergent role of metabolic phenotyping in dynamic patient stratification. *Expert Opin Drug Metab Toxicol* 2014;**10**(7):915–19.

217. Nicholson JK, Holmes E, Kinross JM, *et al*. Metabolic phenotyping in clinical and surgical environments. *Nature* 2012;**491**(7424):384–92.

218. Everett JR, Loo RL, Pullen FS. Pharmacometabonomics and personalized medicine. *Ann Clin Biochem* 2013;**50**(Pt 6):523–45.

219. Nicholson JK, Wilson ID, Lindon JC. Pharmacometabonomics as an effector for personalized medicine. *Pharmacogenomics* 2011;**12**(1):103–11.

220. Everett JR. Pharmacometabonomics in humans: a new tool for personalized medicine. *Pharmacogenomics* 2015;**16**(7):737–54.

221. Elmariah S, Farrell LA, Daher M, *et al*. Metabolite profiles predict acute kidney injury and mortality in patients undergoing transcatheter aortic valve replacement. *J Am Heart Assoc* 2016;**5**:e002712.

222. Barba I, de LG, Martin E, *et al*. Nuclear magnetic resonance-based metabolomics predicts exercise-induced ischemia in patients with suspected coronary artery disease. *Magn Reson Med* 2008;**60**(1):27–32.

223. Wei S, Liu L, Zhang J, *et al*. Metabolomics approach for predicting response to neoadjuvant chemotherapy for breast cancer. *Mol Oncol* 2013;**7**(3):297–307.

224. Andras I, Crisan N, Vesa S, *et al*. Serum metabolomics can predict the outcome of first systematic transrectal prostate biopsy in patients with PSA <10 ng/ml. *Future Oncol* 2017;**13**:1793–800.

225. Yang A, Troup M, Ho JWK. Scalability and validation of big data bioinformatics software. *Comput Struct Biotechnol J* 2017;**15**:379–86.

226. Greene CS, Tan J, Ung M, *et al*. Big data bioinformatics. *J Cell Physiol* 2014;**229**(12):1896–900.

227. Nielson JL, Cooper SR, Yue JK, *et al*. Uncovering precision phenotype-biomarker associations in traumatic brain injury using topological data analysis. *PLoS One* 2017;**12**(3):e0169490.

228. Luo Y, Ahmad FS, Shah SJ. Tensor factorization for precision medicine in heart failure with preserved ejection fraction. *J Cardiovasc Transl Res* 2017;**10**(3):305–12.

229. Mendes P. Emerging bioinformatics for the metabolome. *Brief Bioinform* 2002;**3**(2):134–45.

230. Worley B, Powers R. Multivariate analysis in metabolomics. *Curr Metab* 2013;**1**(1):92–107.

231. Eisen MB, Spellman PT, Brown PO, *et al*. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA* 1998;**95**:14863–8.

232. Odunsi K, Wollman RM, Ambrosone CB, *et al*. Detection of epithelial ovarian cancer using 1H-NMR-based metabonomics. *Int J Cancer* 2005;**113**(5):782–8.

233. Tamayo P, Slonim D, Mesirov J, *et al*. Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. *Proc Natl Acad Sci USA* 1999;**96**(6):2907–12.

234. Heinemann J, Mazurie A, Tokmina-Lukaszewska M, *et al*. Application of support vector machines to metabolomics experiments with limited replicates. *Metabolomics* 2014; **10**(6):1121–8.

235. Mahadevan S, Shah SL, Marrie TJ, *et al*. Analysis of metabolomic data using support vector machines. *Anal Chem* 2008; **80**(19):7562–70.

236. Gromski PS, Muhamadali H, Ellis DI, *et al*. A tutorial review: metabolomics and partial least squares-discriminant analysis–a marriage of convenience or a shotgun wedding. *Anal Chim Acta* 2015;**879**:10–23.

237. MacIntyre DA, Jimenez B, Lewintre EJ, *et al*. Serum metabolome analysis by 1H-NMR reveals differences between chronic lymphocytic leukaemia molecular subgroups. *Leukemia* 2010;**24**(4):788–97.

238. Bhuvaneshwar K, Belouali A, Singh V, *et al*. G-DOC plus—an integrative bioinformatics platform for precision medicine. *BMC Bioinformatics* 2016;**17**(1):193.

239. Li G, Bankhead P, Dunne PD, *et al*. Embracing an integromic approach to tissue biomarker research in cancer: perspectives and lessons learned. *Brief Bioinform* 2017;**18**:634–46.

240. Hintzsche J, Kim J, Yadav V, *et al*. IMPACT: a whole-exome sequencing analysis pipeline for integrating molecular profiles with actionable therapeutics in clinical samples. *J Am Med Inform Assoc* 2016;**23**(4):721–30.

241. Zhang X, Kuivenhoven JA, Groen AK. Forward individualized medicine from personal genomes to interactomes. *Front Physiol* 2015;**6**:364.

242. Quinn RA, Nothias LF, Vining O, *et al*. Molecular networking as a drug discovery, drug metabolism, and precision medicine strategy. *Trends Pharmacol Sci* 2017;**38**(2):143–54.

243. Li S, Todor A, Luo R. Blood transcriptomics and metabolomics for personalized medicine. *Comput Struct Biotechnol J* 2016;**14**:1–7.

244. Smith CA, Want EJ, O'Maille G, *et al*. XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal Chem* 2006;**78**(3):779–87.

245. Benton HP, Wong DM, Trauger SA, *et al*. XCMS(2): processing tandem mass spectrometry data for metabolite identification and structural characterization. *Anal Chem* 2008;**80**(16): 6382–9.

246. Huan T, Forsberg EM, Rinehart D, *et al*. Systems biology guided by XCMS online metabolomics. *Nat Methods* 2017; **14**(5):461–2.

247. Katajamaa M, Miettinen J, Oresic M. MZmine: toolbox for processing and visualization of mass spectrometry based molecular profile data. *Bioinformatics* 2006;**22**(5):634–6.

248. Uppal K, Soltow QA, Strobel FH, *et al*. xMSanalyzer: automated pipeline for improved feature detection and downstream analysis of large-scale, non-targeted metabolomics data. *BMC Bioinformatics* 2013;**14**(1):15.

249. Pfeuffer J, Sachsenberg T, Alka O, *et al*. OpenMS—a platform for reproducible analysis of mass spectrometry data. *J Biotechnol* 2017;**261**:142–8.

250. Xia J, Wishart DS. Using MetaboAnalyst 3.0 for comprehensive metabolomics data analysis. *Curr Protoc Bioinformatics* 2016;**55**:14.10.1–91.

251. Xia J, Sinelnikov IV, Han B, *et al*. MetaboAnalyst 3.0–making metabolomics more meaningful. *Nucleic Acids Res* 2015; **43**(W1):W251–7.

252. Xia J, Mandal R, Sinelnikov IV, *et al*. MetaboAnalyst 2.0–a comprehensive server for metabolomic data analysis. *Nucleic Acids Res* 2012;**40**:W127–33.

253. Jia P, Wang Q, Chen Q, *et al*. MSEA: detection and quantification of mutation hotspots through mutation set enrichment analysis. *Genome Biol* 2014;**15**(10):489.

254. Xia J, Wishart DS. MSEA: a web-based tool to identify biologically meaningful patterns in quantitative metabolomic data. *Nucleic Acids Res* 2010;**38**:W71–7.

255. Teta R, Della Sala G, Glukhov E, *et al*. Combined LC-MS/MS and molecular networking approach reveals new cyanotoxins from the 2014 cyanobacterial bloom in Green Lake, Seattle. *Environ Sci Technol* 2015;**49**(24):14301–10.

256. Xia J, Wishart DS. MetPA: a web-based metabolomics tool for pathway analysis and visualization. *Bioinformatics* 2010; **26**(18):2342–4.

257. Tautenhahn R, Cho K, Uritboonthai W, *et al*. An accelerated workflow for untargeted metabolomics using the METLIN database. *Nat Biotechnol* 2012;**30**(9):826–8.

258. Sana TR, Roark JC, Li X, *et al*. Molecular formula and METLIN personal metabolite database matching applied to the identification of compounds generated by LC/TOF-MS. *J Biomol Tech* 2008;**19**:258–66.

259. Smith CA, O'Maille G, Want EJ, *et al*. METLIN: a metabolite mass spectral database. *Ther Drug Monit* 2005;**27**(6):747–51.

260. Kanehisa M. KEGG bioinformatics resource for plant genomics and metabolomics. *Methods Mol Biol* 2016;**1374**:55–70.

261. Wagele B, Witting M, Schmitt-Kopplin P, *et al*. MassTRIX reloaded: combined analysis and visualization of transcriptome and metabolome data. *PLoS One* 2012;**7**(7):e39860.

262. Suhre K, Schmitt-Kopplin P. MassTRIX: mass translator into pathways. *Nucleic Acids Res* 2008;**36**:W481–4.

263. Cui Q, Lewis IA, Hegeman AD, *et al*. Metabolite identification via the madison metabolomics consortium database. *Nat Biotechnol* 2008;**26**(2):162–4.

264. Southan C, Sitzmann M, Muresan S. Comparing the chemical structure and protein content of ChEMBL, DrugBank, human metabolome database and the therapeutic target database. *Mol Inform* 2013;**32**(11–12):881–97.

265. Sud M, Fahy E, Cotter D, *et al*. LIPID MAPS-nature lipidomics gateway: an online resource for students and educators interested in lipids. *J Chem Educ* 2012;**89**(2):291–2.

266. Sud M, Fahy E, Cotter D, *et al*. LMSD: LIPID MAPS structure database. *Nucleic Acids Res* 2007;**35**:D527–32.

267. Cotter D, Maer A, Guda C, *et al*. LMPD: LIPID MAPS proteome database. *Nucleic Acids Res* 2006;**34**:D507–10.

268. Kim S, Thiessen PA, Bolton EE, *et al*. PubChem substance and compound databases. *Nucleic Acids Res* 2016;**44**(D1): D1202–13.

269. Little JL, Williams AJ, Pshenichnov A, *et al*. Identification of "known unknowns" utilizing accurate mass data and ChemSpider. *J Am Soc Mass Spectrom* 2012;**23**(1):179–85.

270. Contreras AV, Cocom-Chan B, Hernandez-Montes G, *et al*. Host-microbiome interaction and cancer: potential application in precision medicine. *Front Physiol* 2016;**7**:606.

271. Afshinnekoo E, Chou C, Alexander N, *et al*. Precision metage-nomics: rapid metagenomic analyses for infectious disease diagnostics and public health surveillance. *J Biomol Tech* 2017;**28**(1):40–5.

272. Shanahan F. The gut microbiota-a clinical perspective on les-sons learned. *Nat Rev Gastroenterol Hepatol* 2012;**9**(10):609–14.

273. Marcobal A, Kashyap PC, Nelson TA, *et al*. A metabolomic view of how the human gut microbiota impacts the host metabolome using humanized and gnotobiotic mice. *ISME J* 2013;**7**(10):1933–43.

274. Li H, He J, Jia W. The influence of gut microbiota on drug metabolism and toxicity. *Expert Opin Drug Metab Toxicol* 2016; **12**(1):31–40.

275. Smirnov KS, Maier TV, Walker A, *et al*. Challenges of metabo-lomics in human gut microbiota research. *Int J Med Microbiol* 2016;**306**(5):266–79.

276. Hou W, Zhong D, Zhang P, *et al*. A strategy for the targeted metabolomics analysis of 11 gut microbiota-host co-metab-olites in rat serum, urine and feces by ultra high perform-ance liquid chromatography-tandem mass spectrometry. *J Chromatogr A* 2016;**1429**:207–17.

277. Smits SA, Marcobal A, Higginbottom S, *et al*. Individualized responses of gut microbiota to dietary intervention modeled in humanized mice. *mSystems* 2016;**1**(5):1.

278. Zeng Y, Luo L, Hou W, *et al*. Targeted metabolomics analysis of aromatic amino acids and their gut microbiota-host co-metabolites in rat serum and urine by liquid

279. Aguiar-Pulido V, Huang W, Suarez-Ulloa V, *et al*. Metagenomics, metatranscriptomics, and metabolomics approaches for microbiome analysis. *Evol Bioinform Online* 2016;**12**:5–16.

chromatography with tandem mass spectrometry. *J Sep Sci* 2017;**40**(16):3221–30.

280. Bryce AH, Egan JB, Borad MJ, *et al*. Experience with precision genomics and tumor board, indicates frequent target identi-fication, but barriers to delivery. *Oncotarget* 2017;**8**:27145–54.

281. Servant N, Roméjon J, Gestraud P, *et al*. Bioinformatics for precision medicine in oncology: principles and application to the SHIVA clinical trial. *Front Genet* 2014;**5**:152.

282. Roy Choudhury A, Cheng T, Phan L, *et al*. Supporting preci-sion medicine by data mining across multi-disciplines: an integrative approach for generating comprehensive linkages between single nucleotide variants (SNVs) and drug-binding sites. *Bioinformatics* 2017;**33**(11):1621–9.

283. Tenenbaum JD. Translational bioinformatics: past, present, and future. *Genomics Proteomics Bioinformatics* 2016;**14**(1): 31–41.

284. Lee H, Ross JS. The potential role of comprehensive genomic profiling to guide targeted therapy for patients with biliary cancer. *Therap Adv Gastroenterol* 2017;**10**(6):507–20.

285. Malgerud L, Lindberg J, Wirta V, *et al*. Bioinformatory-assisted analysis of next-generation sequencing data for precision medicine in pancreatic cancer. *Mol Oncol* 2017; **11**(10):1413–29.